

2

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/777,317

DATE: 07/25/2001

TIME: 10:45:00

Input Set : N:\Crif3\RULE60\09777317.txt

Output Set: N:\CRF3\07252001\I777317.raw

SEQUENCE LISTING

2 (1) GENERAL INFORMATION:

3 (i) APPLICANT: Gossen, Manfred

4 Bujard, Hermann

5 Salfeld, Jochen

6 Voss, Jeffrey

8 (ii) TITLE OF INVENTION: Methods for Regulating Gene Expression

9 (iii) NUMBER OF SEQUENCES: 10

10 (iv) CORRESPONDENCE ADDRESS:

12 (A) ADDRESSEE: Lahive & Cockfield

13 (B) STREET: 60 State Street, Suite 510

14 (C) CITY: Boston

15 (D) STATE: Massachusetts

16 (E) COUNTRY: USA

17 (F) ZIP: 02109-1875

18 (v) COMPUTER READABLE FORM:

19 (A) MEDIUM TYPE: Floppy disk

20 (B) COMPUTER: IBM PC compatible

21 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

22 (D) SOFTWARE: ASCII text

24 (vi) CURRENT APPLICATION DATA:

C--> 25 (A) APPLICATION NUMBER: US/09/777,317

C--> 26 (B) FILING DATE: 25-Jul-2001

27 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

29 (A) APPLICATION NUMBER: 08/479,306

30 (B) FILING DATE:

32 (A) APPLICATION NUMBER: 08/076,327

33 (B) FILING DATE: 14-JUN-1993

34 (viii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: Giulio A. DeConti, Jr.

36 (B) REGISTRATION NUMBER: 31,503

37 (C) REFERENCE/DOCKET NUMBER: BBI-013CP3

38 (ix) TELECOMMUNICATION INFORMATION:

39 (A) TELEPHONE: (617) 227-7400

40 (B) TELEFAX: (617) 227-5941

42 (2) INFORMATION FOR SEQ ID NO: 1:

44 (i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 1008 base pairs

46 (B) TYPE: nucleic acid

47 (C) STRANDEDNESS: double

48 (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: DNA (genomic)

52 (vi) ORIGINAL SOURCE:

53 (A) ORGANISM: Herpes Simplex Virus

54 (B) STRAIN: K12, KOS

56 (vii) IMMEDIATE SOURCE:

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/777,317

DATE: 07/25/2001

TIME: 10:45:00

Input Set : N:\Crf3\RULE60\09777317.txt

Output Set: N:\CRF3\07252001\I777317.raw

```

57          (B) CLONE: tTA transactivator
59      (ix) FEATURE:
60          (A) NAME/KEY: exon
61          (B) LOCATION: 1..1008
63      (ix) FEATURE:
64          (A) NAME/KEY: mRNA
65          (B) LOCATION: 1..1008
67      (ix) FEATURE:
68          (A) NAME/KEY: misc. binding
69          (B) LOCATION: 1..207
71      (ix) FEATURE:
72          (A) NAME/KEY: misc. binding
73          (B) LOCATION: 208..335
75      (ix) FEATURE:
76          (A) NAME/KEY: CDS
77          (B) LOCATION: 1..1005
79      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
80 ATG TCT AGA TTA GAT AAA AGT AAA GTG ATT AAC AGC GCA TTA GAG CTG      48
81 Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
82   1           5           10           15
84 CTT AAT GAG GTC GGA ATC GAA GGT TTA ACA ACC CGT AAA CTC GCC CAG      96
85 Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
86           20           25           30
88 AAG CTA GGT GTA GAG CAG CCT ACA TTG TAT TGG CAT GTA AAA AAT AAG      144
89 Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys
90           35           40           45
92 CGG GCT TTG CTC GAC GCC TTA GCC ATT GAG ATG TTA GAT AGG CAC CAT      192
93 Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
94           50           55           60
96 ACT CAC TTT TGC CCT TTA GAA GGG GAA AGC TGG CAA GAT TTT TTA CGT      240
97 Thr His Phe Cys Pro Leu Glu Gly Glu Ser Trp Gln Asp Phe Leu Arg
98   65           70           75           80
100 AAT AAG GCT AAA AGT TTT AGA TGT GCT TTA CTA AGT CAT CGC GAT GGA      288
101 Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asp Gly
102           85           90           95
104 GCA AAA GTA CAT TTA GGT ACA CGG CCT ACA GAA AAA CAG TAT GAA ACT      336
105 Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr
106           100          105          110
108 CTC GAA AAT CAA TTA GCC TTT TTA TGC CAA CAA GGT TTT TCA CTA GAG      384
109 Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu
110           115          120          125
112 AAT GCA TTA TAT GCA CTC AGC GCT GTG GGG CAT TTT ACT TTA GGT TGC      432
113 Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys
114           130          135          140
116 GTA TTG GAA GAT CAA GAG CAT CAA GTC GCT AAA GAA GAA AGG GAA ACA      480
117 Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr
118 145          150          155          160
120 CCT ACT ACT GAT AGT ATG CCG CCA TTA TTA CGA CAA GCT ATC GAA TTA      528
121 Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/777,317

DATE: 07/25/2001

TIME: 10:45:00

Input Set : N:\Crf3\RULE60\09777317.txt

Output Set: N:\CRF3\07252001\I777317.raw

```

122          165          170          175
124 TTT GAT CAC CAA GGT GCA GAG CCA GCC TTC TTA TTC GGC CTT GAA TTG      576
125 Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu
126          180          185          190
128 ATC ATA TGC GGA TTA GAA AAA CAA CTT AAA TGT GAA AGT GGG TCC GCG      624
129 Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala
130          195          200          205
132 TAC AGC CGC GCG CGT ACG AAA AAC AAT TAC GGG TCT ACC ATC GAG GGC      672
133 Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly
134          210          215          220
136 CTG CTC GAT CTC CCG GAC GAC GAC GCC CCC GAA GAG GCG GGG CTG GCG      720
137 Leu Leu Asp Leu Pro Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala
138 225          230          235          240
140 GCT CCG CGC CTG TCC TTT CTC CCC GCG GGA CAC ACG CGC AGA CTG TCG      768
141 Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser
142          245          250          255
144 ACG GCC CCC CCG ACC GAT GTC AGC CTG GGG GAC GAG CTC CAC TTA GAC      816
145 Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp
146          260          265          270
148 GGC GAG GAC GTG GCG ATG GCG CAT GCC GAC GCG CTA GAC GAT TTC GAT      864
149 Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp
150          275          280          285
152 CTG GAC ATG TTG GGG GAC GGG GAT TCC CCG GGT CCG GGA TTT ACC CCC      912
153 Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro
154          290          295          300
156 CAC GAC TCC GCC CCC TAC GGC GCT CTG GAT ATG GCC GAC TTC GAG TTT      960
157 His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe
158 305          310          315          320
160 GAG CAG ATG TTT ACC GAT CCC CTT GGA ATT GAC GAG TAC GGT GGG TAG      1008
161 Glu Gln Met Phe Thr Asp Pro Leu Gly Ile Asp Glu Tyr Gly Gly
162          325          330          335
164 (2) INFORMATION FOR SEQ ID NO: 2:
165     (i) SEQUENCE CHARACTERISTICS:
166         (A) LENGTH: 335 amino acids
167         (B) TYPE: amino acid
168         (D) TOPOLOGY: linear
169     (ii) MOLECULE TYPE: protein
170     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
171 Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
172  1          5          10          15
174 Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
175          20          25          30
177 Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys
178          35          40          45
180 Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
181          50          55          60
183 Thr His Phe Cys Pro Leu Gly Glu Ser Trp Gln Asp Phe Leu Arg
184 65          70          75          80
186 Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asp Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/777,317

DATE: 07/25/2001

TIME: 10:45:00

Input Set : N:\Crf3\RULE60\09777317.txt

Output Set: N:\CRF3\07252001\I777317.raw

```

187           85           90           95
189 Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr
190           100           105           110
192 Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu
193           115           120           125
195 Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys
196           130           135           140
198 Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr
199 145           150           155           160
201 Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu
202           165           170           175
204 Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu
205           180           185           190
207 Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala
208           195           200           205
210 Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly
211           210           215           220
213 Leu Leu Asp Leu Pro Asp Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala
214 225           230           235           240
216 Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser
217           245           250           255
219 Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp
220           260           265           270
222 Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp
223           275           280           285
225 Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro
226           290           295           300
228 His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe
229 305           310           315           320
231 Glu Gln Met Phe Thr Asp Pro Leu Gly Ile Asp Glu Tyr Gly Gly
232           325           330           335

```

235 (2) INFORMATION FOR SEQ ID NO: 3:

236 (i) SEQUENCE CHARACTERISTICS:

237 (A) LENGTH: 894 base pairs

238 (B) TYPE: nucleic acid

239 (C) STRANDEDNESS: double

240 (D) TOPOLOGY: linear

242 (ii) MOLECULE TYPE: DNA (genomic)

244 (vi) ORIGINAL SOURCE:

245 (A) ORGANISM: Herpes Simplex Virus

246 (B) STRAIN: K12, KOS

247 (C) INDIVIDUAL ISOLATE: tTAS transactivator

249 (ix) FEATURE:

250 (A) NAME/KEY: exon

251 (B) LOCATION: 1..894

253 (ix) FEATURE:

254 (A) NAME/KEY: mRNA

255 (B) LOCATION: 1..894

257 (ix) FEATURE:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/777,317

DATE: 07/25/2001

TIME: 10:45:00

Input Set : N:\Crif3\RULE60\09777317.txt

Output Set: N:\CRF3\07252001\I777317.raw

```

258      (A) NAME/KEY: misc. binding
259      (B) LOCATION: 1..207
261      (ix) FEATURE:
262      (A) NAME/KEY: misc. binding
263      (B) LOCATION: 208..297
265      (ix) FEATURE:
266      (A) NAME/KEY: CDS
267      (B) LOCATION: 1..891
269      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
270 ATG TCT AGA TTA GAT AAA AGT AAA GTG ATT AAC AGC GCA TTA GAG CTG      48
271 Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
272 1 5 10 15
273 CTT AAT GAG GTC GGA ATC GAA GGT TTA ACA ACC CGT AAA CTC GCC CAG      96
274 Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
275 20 25 30
276 AAG CTA GGT GTA GAG CAG CCT ACA TTG TAT TGG CAT GTA AAA AAT AAG      144
277 Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys
278 35 40 45
279 CGG GCT TTG CTC GAC GCC TTA GCC ATT GAG ATG TTA GAT AGG CAC CAT      192
280 Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
281 50 55 60
282 ACT CAC TTT TGC CCT TTA GAA GGG GAA AGC TGG CAA GAT TTT TTA CGT      240
283 Thr His Phe Cys Pro Leu Glu Gly Glu Ser Trp Gln Asp Phe Leu Arg
284 65 70 75 80
285 AAT AAC GCT AAA AGT TTT AGA TGT GCT TTA CTA AGT CAT CGC GAT GGA      288
286 Asn Asn Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asp Gly
287 85 90 95
288 GCA AAA GTA CAT TTA GGT ACA CGG CCT ACA GAA AAA CAG TAT GAA ACT      336
289 Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr
290 100 105 110
291 CTC GAA AAT CAA TTA GCC TTT TTA TGC CAA CAA GGT TTT TCA CTA GAG      384
292 Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu
293 115 120 125
294 AAT GCA TTA TAT GCA CTC AGC GCT GTG GGG CAT TTT ACT TTA GGT TGC      432
295 Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys
296 130 135 140
297 GTA TTG GAA GAT CAA GAG CAT CAA GTC GCT AAA GAA GAA AGG GAA ACA      480
298 Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr
299 145 150 155 160
300 CCT ACT ACT GAT AGT ATG CCG CCA TTA TTA CGA CAA GCT ATC GAA TTA      528
301 Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu
302 165 170 175
303 TTT GAT CAC CAA GGT GCA GAG CCA GCC TTC TTA TTC GGC CTT GAA TTG      576
304 Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu
305 180 185 190
306 ATC ATA TGC GGA TTA GAA AAA CAA CTT AAA TGT GAA AGT GGG TCT GAT      624
307 Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Asp
308 195 200 205
309 CCA TCG ATA CAC ACG CGC AGA CTG TCG ACG GCC CCC CCG ACC GAT GTC      672

```

VERIFICATION SUMMARY

DATE: 07/25/2001

PATENT APPLICATION: US/09/777,317

TIME: 10:45:01

Input Set : N:\Crf3\RULE60\09777317.txt

Output Set: N:\CRF3\07252001\I777317.raw

L:25 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:26 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

STATISTICS SUMMARY

PATENT APPLICATION: US/09/777,317

DATE: 07/25/2001

TIME: 10:45:01

Input Set : N:\Crif3\RULE60\09777317.txt

Output Set: N:\CRF3\07252001\I777317.raw

Application Serial Number: US/09/777,317

Alpha or Numeric: Alpha

Application Class:

Application File Date: 07-25-2001

Art Unit:

Software Application: Other

Total Number of Sequences: 10

Total Nucleotides: 14449

Total Amino Acids: 632

Number of Errors: 0

Number of Warnings: 0

Number of Corrections: 1

MESSAGE SUMMARY

220 C: 1 (Keyword misspelled or invalid format)